



UWS Academic Portal

A QoE and simulator sickness evaluation of a smart-exercise-bike virtual reality system via user feedback and physiological signals

Katsigiannis, Stamos; Willis, Rhys; Ramzan, Naeem

Published in:
IEEE Transactions on Consumer Electronics

DOI:
[10.1109/TCE.2018.2879065](https://doi.org/10.1109/TCE.2018.2879065)

Published: 28/02/2019

Document Version
Peer reviewed version

[Link to publication on the UWS Academic Portal](#)

Citation for published version (APA):
Katsigiannis, S., Willis, R., & Ramzan, N. (2019). A QoE and simulator sickness evaluation of a smart-exercise-bike virtual reality system via user feedback and physiological signals. *IEEE Transactions on Consumer Electronics*, 65(1), 119-127. <https://doi.org/10.1109/TCE.2018.2879065>

General rights

Copyright and moral rights for the publications made accessible in the UWS Academic Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact pure@uws.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

“© © 2018 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

A QoE and Simulator Sickness Evaluation of a Smart-Exercise-Bike Virtual Reality System via User Feedback and Physiological Signals

Stamos Katsigiannis, Rhys Willis, and Naeem Ramzan, *Senior Member, IEEE*

Abstract—In this work, we proposed, designed, and evaluated a smart-exercise-bike virtual reality system. The proposed system uses a custom microcontroller-based board in order to transmit signals from a stationary exercise bike to a computer, while a commercially available head mounted display (HMD) is used in order to immerse the users into a virtual reality environment. Various quality settings related to texture resolution and to the frame rate of the rendered environment were evaluated by the users, while simulator sickness was assessed using the Simulator Sickness Questionnaire. Furthermore, physiological responses in terms of electrocardiography (ECG) and galvanic skin response (GSR) signals were captured for the whole duration of the experiments in order to better assess the effects of rendering quality and simulator sickness on the quality of experience (QoE) of the proposed smart-exercise-bike virtual reality system. The effect and correlation between rendering quality and simulator sickness is then discussed in this work.

Index Terms—smart-bike, cycling exergame, exergaming, virtual reality, HMD, QoE, simulator sickness, physiological signals, ECG, GSR

I. INTRODUCTION

VIRTUAL REALITY (VR) devices are currently flooding the technology market. While VR technology's most mainstream focus is the gaming and entertainment industries, it has also found applications within manufacturing, architecture, health care, and physical exercise [1]–[4]. Since its early days, VR has been used as a medium to train professionals in a variety of disciplines including medicine, the military, and fire fighting [5]. Furthermore, the benefits of VR in training and education environments have been demonstrated in various studies [6]. In recent years, the addition of gaming elements to VR training applications has led to the emergence of *serious games* that have the potential to make the training environments more engaging to the general public [7]. With commercial VR equipment (Head-Mounted Displays - HMDs) becoming available to the public at affordable prices [8], VR technology has the potential to be integrated into various everyday applications.

One field that has significantly benefited from the availability of VR technology is *exergaming* [9], [10]. Exergaming refers to video games that include some form of physical exercise and rely on technology that can track body movement

[11]. Among various available exergames, cycling exergames have attracted a lot of attention due to facilitating the use of stationary exercise bikes by providing audiovisual applications that increase the motivation levels of the users [12]. The ability to promote physical activity is a highly advantageous characteristic of exergames [13]. According to Lee et al. [14], physical inactivity is the fourth leading cause of death globally and leads to an increased risk of developing death-threatening diseases, such as cardiovascular disease and type 2 diabetes [15]. While 75 to 150 minutes of exercise per week are recommended for leading a healthy lifestyle [16], 40 to 65% of people that start a physical exercise regimen are expected to drop out within 3 to 6 months [17]. As a result, the use of equipment and methods that may motivate people to exercise more would be highly beneficial to the individual, as well as to society due to the potential reduction in health care expenses.

Designing a VR cycling exergame is a challenging task. Apart from the hardware challenges, the software, especially the VR application, has to be able to facilitate user engagement and keep the users comfortable and motivated to continue using it. According to Hagen et al. [18], in order for a cycling exergame to be both engaging and beneficial to the health of the user, it has to achieve a precise balance between interaction design and exercise physiology.

With VR in such increasingly high demand across a broad spectrum of industries and with a large variety of uses, it becomes necessary to have a better understanding of how the quality of VR content can directly affect user experience. It is well documented that HMDs often provoke discomfort and even nausea [19], a condition known as *simulator sickness*. Being a sub-category of motion-sickness, simulator sickness arises due to sensory conflict between what the users see and what they feel [20]. Simulator sickness can be detrimental to the ability of a VR cycling exergame to attract users and motivate them to use it. Consequently, it is important to study the effects of simulator sickness, as well as of other quality parameters, to the perceived Quality of Experience (QoE) of the users.

In this work, we propose and evaluate a physical exercise system comprised of a smart-exercise-bike and a virtual reality platform that allows the users to engage into physical exercise while immersed within a virtual environment. The smart-exercise-bike system is designed as a generic kit that can be attached to any stationary exercise bike in order to provide “smart” capabilities and is able to operate either as a generic gaming controller, or in combination with dedicated

Manuscript received Month X, 2018;

S. Katsigiannis, R. Willis, and N. Ramzan were with the School of Computing, Engineering & Physical Sciences, University of the West of Scotland, Paisley, PA1 2BE, UK. e-mail: {Stamos.Katsigiannis, Rhys.Willis, Naeem.Ramzan}@uws.ac.uk.

Digital object identifier ...

applications, such as the VR cycling application discussed in this work. The proposed system allows the users to control a virtual bike by pedalling a physical stationary exercise bike which transmits the pedalling signals to a computer. The users would see the bike moving around the virtual world by wearing an HMD. Targeting a wide spectrum of casual users, the proposed system must be evaluated in terms of perceptual visual quality and general quality of experience (QoE). Furthermore, it is important to study whether and to what extent the use of the proposed system will provoke symptoms of simulator sickness. To this end, we conducted trials where users were asked to use the examined system under various pre-defined quality settings and to subsequently report about visual quality and simulator sickness symptoms. Physiological responses in terms of electrocardiography (ECG) and galvanic skin response (GSR) signals were captured for the whole duration of the experiments in order to better assess the effects of visual quality and simulator sickness on the quality of experience of the system, since modalities such as the ECG and GSR have been shown to correlate with the perceived visual quality [21], with stress levels [22], and immersion [23].

The rest of this paper is organised in five sections. Related work is described in Section II, while the hardware and software design of the proposed system is described in Section III. Section IV provides a description of the experimental procedure and results are presented and discussed in Section V. Finally, conclusions are drawn in Section VI.

II. BACKGROUND

Interactive video game exercise has been shown to be effective in enhancing exercise adherence, as well as various health status markers [24]. The availability and low cost of stationary exercise bikes and computers, makes cycling exergames an ideal platform for indoors training at the users' homes. Cycling exergames have been shown to increase the motivation of their users to engage into physical exercise [9], [12], [25]. Göbel et al. [26] presented a set of simple video games controlled by a stationary bike, aiming at improving the long-term motivation of users to engage into physical exercise. Hoda et al. [27] examined the acceptance rates of users when given the opportunity to use a cycling exergame system. Both their quantitative and qualitative results showed that exergames are more entertaining and more encouraging to people with basic activity levels.

In another work, Silva and El Saddik [28] examined the use of a smart-phone as an input controller for a cycling exergame, as well as other types of exergames, showing that smart-phones can be suitable controllers for such games. Their approach offers support for the design of configurable and flexible exertion interfaces that can be integrated with pre-existing games without the need for any modification. Müller et al. [12] designed a cycling game controller and developed a 3D first person cycling game designed to provoke emotions with game elements in different game settings, like timed race, parcours traversal, and virtual world exploration. Their study showed that crafted computer exergame elements are able

to provoke emotions to the users, thus by following specific design patterns the games can facilitate and boost their users' engagement with them.

With the recent availability of VR technology, cycling exergames can benefit from the introduction of VR by providing a more immerse and real-life experience that could potentially further increase the motivation of their users for engaging with them. Bruun-Pedersen et al. [25] examined the combination of conventional chair-based cycling exercise with four recreational virtual environments in order to increase the intrinsic motivation of a group of nursing home residents. Comparisons to a period before the availability of the proposed system showed that it significantly increased the motivation of the nursing home residents to keep using the augmented bike for physical exercise. Grani and Bruun-Pedersen [29] proposed Giro, a wireless tracker for improving functionality and experience of VR cycling systems. Their system was based on a manuped, a stationary cycling device that allows both feet and hand pedalling in order to accommodate the rehabilitation of patients with various conditions. The enhanced accuracy of the proposed system's tracking was demonstrated to be beneficial for VR rehabilitation in physical therapy, upholding the requirements of the medical partners. Barathi et al. [9] studied the use of VR in a cycling-based exergame in order to improve performance while maintaining intrinsic motivation in high intensity VR exergaming. In their work, the authors proposed and evaluated "interactive feedforward", a model of self competition against an improved self model of the player, for rapidly improving performance in high-intensity interval training cycling VR exergames, demonstrating promising potential for increasing performance and motivation.

Apart from research works, commercial cycling exergame products have recently appeared in the market, utilising both VR and more traditional gaming approaches.

III. HARDWARE & SOFTWARE DESIGN

The proposed smart-exercise-bike VR system consists of three components: (a) the bike module, (b) the bike server application, and (c) the VR application. The bike module is used in order to transmit the input from the exercise bike to the computer, where the bike server converts the received data to the required actions for use in the VR application. The bike module, accompanied with the server application, is designed as a "plug and play" kit that can be attached to generic exercise bikes and convert them to gaming controllers, thus transforming generic computer games to cycling exergames. Furthermore, dedicated applications, such as the discussed VR application, can be developed in order to better exploit the system's capabilities and provide enhanced cycling exergaming experiences. An outline of the system and the interaction between its components is given in Fig. 1.

A. The Smart-Exercise-Bike Module

In order to create a truly immersive virtual reality cycling application, a real stationary exercise bike was used as the base of the smart-exercise-bike system. To this end, a custom micro-controller-based circuit was designed and attached to the bike

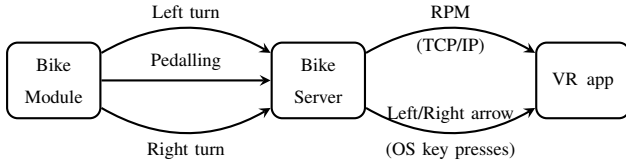


Fig. 1: Outline of the proposed Smart-Exercise-Bike system.

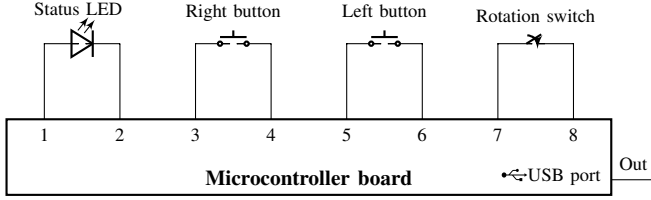


Fig. 2: Basic outline of the bike module. Exact connection and other electronic components have been omitted for simplicity.

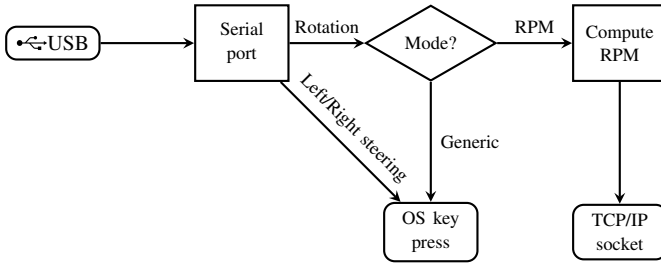


Fig. 3: Outline of bike server application.

in order to detect the rotation of the exercise bike's pedals, provide buttons for steering, and transmit this information to an attached computer in real-time. The bike module was designed as a generic peripheral that can be attached to any exercise bike and connect to any USB enabled computer that can support the accompanying software. A commercially available microcontroller board was used as the basis of the bike module and a custom designed circuit board was used in order to connect all the electronic components to the board.

To detect the rotation of the bike's pedals, a magnetic switch was used, with one part of the switch attached to the bike's body and one on the pedal. At the end of a full rotation of the pedals, the proximity of the magnetic switch's parts causes the circuit to close and the microcontroller then detects the event and transmits a signal to the computer via the USB connection. Support for steering was implemented via two buttons, on the left and the right side of the bike's handles respectively. Upon pressing the button, the microcontroller would transmit to the computer the signal to turn left or right via the USB connection. Continuous pressing of the steering controls was also supported in order to allow for smoother and more natural steering control. An LED was also attached to the board in order to indicate whether user input is registered by the board by blinking when a control signal is received. Furthermore, since the prototype board is not used to power any external components, the board is directly powered by the USB connection used for communication with the computer. An outline of the designed bike module is provided in Fig. 2.

B. The Smart-Exercise-Bike Server

The signals transmitted to the computer by the bike module are handled by a server application that converts them to an appropriate form for use in various applications. When the server application receives a steering signal, it sends a corresponding key press signal to the operating system. For steering left, the server issues a "left arrow" key press, while a "right arrow" key press is issued accordingly for steering right. For single steering signals, the key press signals are issued as "press and release" signals, while for continuous pressing of the steering button, the key press signals are issued as "press and hold" and a "release" signal is issued when the steering button is released. Mapping the input from the bike module to operating system's key presses allows for straightforward integration with any application that supports control using the arrow keys. Furthermore, key mappings can easily change to allow support for various applications.

For the pedalling signals, the server application supports two modes of operation: (a) The generic mode handles each rotation signal as a key press of the "forward arrow" key, similar to the steering control. (b) In the RPM (rotations per minute) mode, used in the examined VR application, the server receives the rotation signals and calculates the number of rotations per minute. Applications can then connect to the server via a local TCP/IP socket and receive the RPM speed in real time. This approach allows multiple applications to connect to the bike server and receive pedalling information, thus providing great flexibility in developing applications for the smart-exercise-bike system. An outline of the bike server application is illustrated in Fig. 3.

C. The Virtual Reality Environment

A simple free roaming environment was developed, consisting of a track that extends within an open area with trees, hills, houses, and large rocks. Objects expected to be present in such an environment (e.g. sign posts, benches, street lights) were also present in order to increase the diversity, the complexity, and the realness of the virtual environment. Furthermore, ambient sounds consistent with the scenery were also played within the game environment. While free roaming, it was also important to limit the playable area to the track. This was achieved by creating a boundary around the track, made up of fences, large rocks and boulders. The track is designed as a closed loop without any dead ends. As a result, the user can roam through the virtual environment for an unlimited distance and time. Fig. 4a shows an aerial view of the map, while a screenshot from first person perspective is shown in Fig. 4b. The speed of the virtual bike is decided by adjusting the rotation speed transmitted from the bike server according to the size of the virtual bike's wheels, thus resulting into more realistic motion. Steering is achieved by setting a minimum turning angle which is gradually increased up to a maximum angle while a steering button is pressed. Releasing the steering button results into the virtual steering handlebar returning to its neutral position (0° turning angle).

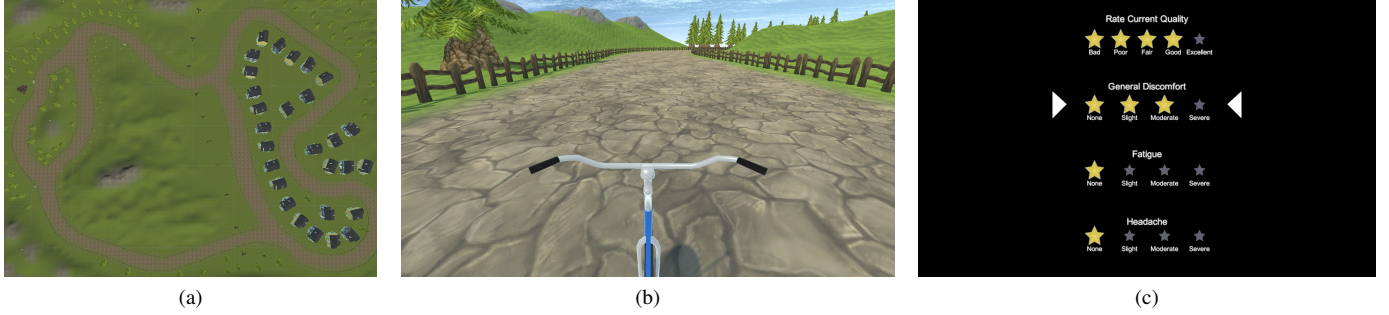


Fig. 4: Screenshots from the virtual environment. (a) Aerial view of the map. (b) First person view. (c) The questionnaire.

IV. EXPERIMENTAL EVALUATION

A. Experiment Setup and Equipment

The participants of this study were asked to use the smart-exercise-bike system while using a commercially available HMD in order to roam through the developed virtual reality environment. The utilised HMD is a virtual reality headset that uses two OLED panels, one for each eye, each having a resolution of 1080×1200 pixels and a 90 Hz refresh rate, providing a total resolution of 2160×1200 pixels and global refresh. It provides a 110° field of view and is also equipped with headphones in order to provide a complete audiovisual experience. The scene at the field of view of the user was continuously mirrored in a 55" TV monitor in order for the supervising researcher to be able to oversee the behaviour of the user and the application. The HMD was connected to a computer equipped with a 4.2 GHz 4-core CPU, 64 GB of RAM, and a high-end graphics card with 12 GB of memory.

While using the smart-exercise-bike system, a wireless 4-lead electrocardiography (ECG) sensor [30] was attached to the participants in order to obtain an ECG signal with a sampling frequency of 512 Hz. ECG is the process of recording the electrical activity of the heart in relation to time. ECG signals have been extensively used for affect recognition [31], [32], and more recently for perceptual video quality evaluation [21]. Furthermore, a Galvanic Skin Response sensor (GSR - Sweating) [33] was also used by attaching two electrodes to the middle and index finger of the right hand of the participant. The sampling rate of the GSR sensor was approximately 20 Hz. GSR provides a measure of the resistance of the skin, which decreases when perspiration increases, an event that usually occurs when a human experiences emotions such as stress or surprise [34].

Both signal sensors were attached to a laptop computer (2.3 GHz 2-core CPU, 4 GB RAM) that was used for signal recording and for plotting the two signals on the screen in real-time. Although the real-time plotting of the acquired signals was not necessary, it facilitated the monitoring of their quality, giving the supervising researcher the opportunity to intervene in case of erroneous signal reception, e.g. if an electrode had been de-attached. Each participant had to perform the experiment once and after each session, the HMD's cushion and lenses, as well as the electrode cables of the ECG and GSR sensors were cleaned with alcohol for hygienic reasons.

B. Experimental scenarios

During the experiment, the participants roam through the virtual environment under various visual quality settings. At start-up, the application begins with a title screen that allows the user to get comfortable wearing the HMD and being immersed into a virtual environment. This also provides the supervising researcher with the opportunity to set-up the experiment by typing the participants name and selecting the duration of the baseline quality setting (30 sec) and of each test stimulus (60 sec). At each quality setting test, the users would first pedal around the virtual environment at the baseline high quality setting (HH) for 30 sec in order to allow time for any effects of the previous test setting to disappear. Then, they had 60 sec to pedal around the virtual environment at the test quality setting and then they had to fill a questionnaire that was presented to them within the virtual environment, as shown in Fig. 4c.

As each of the users would be wearing a VR headset as well as riding an exercise bike, it wasn't feasible for the users to fill out physical surveys. As a result, alternate methods of data collection were experimented with and a digital survey was developed. By using simple horizontal and vertical axis inputs from a gamepad provided by the supervising researcher, the user could navigate through the questions by pushing up and down whilst selecting their answers using left and right. All the answers, along with timestamps about each action in the virtual environment were stored in a log file for further analysis.

During the experiment, each participant had to rate two times each of 6 different quality settings of the virtual reality environment. The variables that were adjusted in order to create the different quality levels were texture resolution and frame rate. For the baseline high quality setting, texture resolution was set to 1024×1024 pixels and the frame rate to 60 fps (HH). Test settings included half (512×512) and a quarter (256×256) of the texture resolution at 60 fps (MH and LH respectively), high texture resolution (1024×1024) at half (30 fps) and at a quarter (15 fps) of the set high frame rate (HM and HL respectively), and random low resolution textures (256×256) at 60 fps (RH). For the random low resolution textures setting (RH), the textures within the environment were set to the highest resolution and random textures would change to low resolution while the user roamed through the

TABLE I: Examined Quality Settings

ID	Quality Setting		Texture resolution (Width × Height)	Frame rate (fps)
	Texture quality	Frame rate		
HH	High	High	1024 × 1024	60
MH	Medium	High	512 × 512	60
LH	Low	High	256 × 256	60
RH	Random low quality textures	High	Random textures at 256 × 256	60
HM	High	Medium	1024 × 1024	30
HL	High	Low	1024 × 1024	15

environment. As a result, various random objects within the virtual environment would appear with lower quality, allowing us to examine whether the users would be able to notice them, thus affecting the overall quality ratings. The quality settings tested in the experiment are outlined in Table I.

C. Experimental procedure

Prior to each experiment, participants were given written instructions explaining their tasks. After reading the instructions, the researcher supervising the experiment provided any further explanations and clarifications required. The participants then proceeded to fill in and sign a consent form, as well as answer a questionnaire about their prior experience with virtual reality applications, with riding bikes, and with the use of exercise bikes. Afterwards, the electrodes for the ECG and GSR sensors were attached to their bodies. After validating the quality of the captured ECG and GSR signals through visual inspection of the plotted signals on the monitoring screen, the HMD was positioned on the head of the participants and adjusted according to the head size. Three random test scenarios were then run in order to familiarise the participants with the test procedure.

The visual quality of the virtual environment was evaluated using *Absolute Category Ratings* (ACR), as defined by ITU-T Rec. P.910 [35]. Under ACR test conditions, the test stimulus is presented once, followed by its rating by the participants. The quality settings shown in Table I were used as test stimuli and participants were asked to rate them using a five-level scale: (1) Bad, (2) Poor, (3) Fair, (4) Good, and (5) Excellent. In order to examine the consistency and the reliability of the results, each test stimulus was presented twice to the participants in random order, ensuring that the same stimulus would never appear twice in a row. Furthermore, to avoid any presentation order effects, the order of the test stimuli shown to each participant was randomised.

Apart from the quality rating, after each stimulus the participants were also asked to evaluate any symptom of simulator sickness that they might have felt. The *Simulator Sickness Questionnaire* (SSQ) [20] was used in order to evaluate the levels of simulator sickness reported by the participants. The SSQ is based on the older *Pensacola Motion Sickness Questionnaire* (MSQ) which contained 28 symptoms associated with or premonitory of motion sickness [36]. From the 28 symptoms of the MSQ, the SSQ utilises only 16 which are

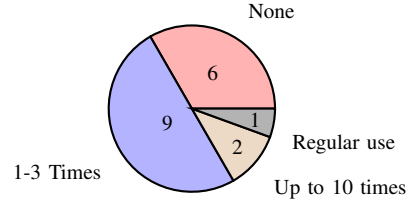


Fig. 5: Participants' experience with virtual reality equipment.

more suitable for simulator sickness. These symptoms are: 1) general discomfort, 2) fatigue, 3) headache, 4) eye strain, 5) difficulty focusing, 6) increased salivation, 7) sweating, 8) nausea, 9) difficulty concentrating, 10) fullness of head, 11) blurred vision, 12) dizziness (eyes open), 13) dizziness (eyes closed), 14) vertigo, 15) stomach awareness, 16) burping, and are clustered into three distinct categories, i.e. *nausea* (N), *oculomotor* (O), and *disorientation* (D). The severity of each of these symptoms was reported by the participants using a 4-point scale (0, 1, 2, 3), with symptom severity regarded as 0: None, 1: Slight, 2: Moderate, and 3: Severe. Weighting of the scores was then applied in order to compute the score for each category (N, O, D), as well as the simulator sickness *Total Score* (TS), as defined in [20].

ECG and GSR recordings were captured from each participant for the whole duration of the experiment. Unix timestamps with millisecond accuracy were used in order to synchronise the ECG and GSR recordings, as well as the test scenarios experienced by the users. After the end of the experimental phase of the study, the captured timestamps were used in order to divide the ECG and GSR recordings into segments referring to each stimulus, in order to examine possible relations between the visual quality or the simulator sickness to the physiological responses.

D. Participants

18 subjects participated in this QoE and simulator sickness evaluation experiment. From these 18, 16 were male and 2 were female, with their age varying between 17 and 36 years ($\mu = 26$, $\sigma = 5.65$). The subjects were undergraduate, postgraduate, and PhD students recruited from the University of the West of Scotland and all had normal or corrected eyesight, without any problem with their vision and depth perception. The subject's visual acuity was tested using Snellen charts, while Ishihara charts were used for the examination of colour vision. Fig. 5 provides information about their prior experience with virtual reality applications, showing that 66.66% of the participants had prior experience with virtual reality technology, while 50% had prior experience with the HMD used in this study. Furthermore, 17 out of 18 participants had ridden a regular bike before and 16 out of 18 had used an exercise bike. Prior experience with either regular or exercise bikes was a requirement for participation in the study. Regarding the levels of physical exercise, 61.11% of the participants reported engaging into physical exercise at least 1-2 times per week, while the rest reported typical daily walking as their only physical exercise.

V. RESULTS AND DISCUSSION

A. QoE

Mean Opinion Scores (MOS) and the respective 95% confidence intervals (CIs) were computed using the participants' quality ratings in order to determine how the participants perceived the visual quality of the virtual environment. As shown in Fig. 6, it is evident that the MOS ratings acquired for each quality setting are all very close to a MOS of 3 ($MOS \in [2.60, 3.47]$), indicating that the users perceived all quality settings to be of "Fair" visual quality in general. As expected, texture resolution was the most significant indicator of the perceived quality. The quality setting using the low resolution textures and high frame rate (LH) received the lowest MOS (2.60) in this study, with the quality setting using the medium resolution textures and high frame rate (MH) receiving a MOS of 2.88, and the one using the high resolution textures and high frame rate (HH) achieving a MOS of 3.19. All the quality settings using high resolution textures received higher MOS than the quality settings with medium and low resolution textures.

Interestingly, the quality setting that used high quality textures with random textures being of lower resolution (RH) achieved an almost similar MOS to the high quality setting HH (3.22 vs 3.19), indicating that the participants did not notice the low quality objects within the virtual environment. It seems that due to the moving speed of the virtual bike and the focus of the user into following the road and not crashing into the fences on the sides, resulted in the users paying less attention to individual objects within the environment. As a result, random low quality tiles on the road, or fences, or trees did not affect the perceived visual quality of the environment. Furthermore, low frame rate had less effect on visual quality than low resolution.

The quality setting with high resolution textures and low frame rate HL received a 3.03 MOS, higher than the one achieved for the low and medium resolution settings with high frame rate (2.88 for MH and 2.60 for LH), but lower than the one achieved for the quality setting with high resolution textures and high frame rate HH (3.19). The most notable outcome of the results analysis was that the quality setting with high texture resolution and medium frame rate (HM) received the highest MOS (3.47) among all the examined quality settings.

To evaluate this unexpected finding, the acquired quality ratings for HH were compared to the quality ratings for HM using a 1-way analysis of variance (ANOVA). Results showed that there is no statistically significant difference between the quality ratings for HH and HM ($p = 0.214$). A non-parametric Wilcoxon Rank Sum Test was also performed between the HH and HM quality ratings, with the null hypothesis that the two distributions are the same, and hence have the same median, but the null hypothesis was not rejected ($p = 0.181$). This finding is further reinforced but the evident overlap between the MOS ratings for the two settings when confidence intervals are taken into consideration (3.19 ± 0.26 for HH vs 3.47 ± 0.30 for HM). Consequently, a statistically significant difference between the quality ratings for HH and HM cannot

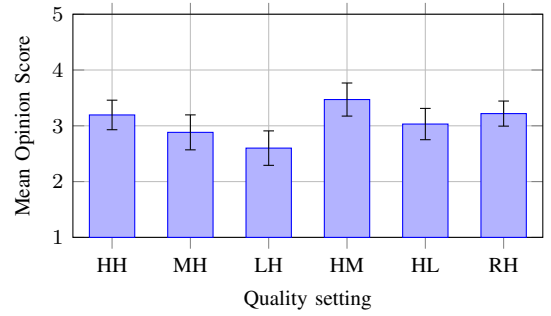


Fig. 6: MOS and 95% CIs obtained for each quality setting.

TABLE II: Within-subjects ANOVA on quality ratings

Factor	Sum Sq	df	Mean Sq	F	p
Texture resolution	12.42	3	4.14	5.57	0.0012
Frame rate	6.84	2	3.42	4.46	0.0127
Quality setting	15.69	5	3.14	4.28	0.0010

Sum Sq: Sum of squares, df: Degrees of freedom, Mean Sq: Mean squared error

be established, indicating that the users were not able to differentiate between the two quality settings. Similarly to the RH setting, it seems that the moving speed of the virtual bike and the focus of the user into following the road and not crashing into the fences, prevented the users from noticing the reduced frame rate, especially in the HH vs HM case (60 fps vs 30 fps). In the HL case (15 fps), the reduction in frame rate was more noticeable, thus the lower MOS, but still not sufficient to severely affect the perception of quality.

A within-subjects analysis of variance (ANOVA) was conducted in order to provide a statistical evaluation of the impact of the texture resolution, the frame rate, and the quality setting (Texture Resolution & Frame rate) on the participants' quality ratings. The ANOVA results on Table II indicate that texture resolution, frame rate, as well as their combination have a statistically significant contribution ($p < 0.013$) to the participants' quality ratings.

B. Simulator Sickness Measurement

The average scores for nausea (N), oculomotor (O), disorientation (D), and the Total Score (TS) acquired from the participants for each quality setting are shown in Fig. 7. From this figure, it is evident that all quality settings led to an average TS of more than or equal to 55 ($\mu_{TS} \geq 55$), with the low resolution quality setting (LH) resulting in the highest average TS (66.14), and the high resolution with medium frame rate setting (HM) resulting in the lowest average TS (55). While the average TS for all participants and quality settings was between 55 and 66.14, the maximum and average TS observed for each individual participant for all the quality settings exhibit high variability, as shown in Fig. 8. The lowest TS was observed for participants #15 ($MaxTS_{\#15} = 3.74$, $AvgTS_{\#15} = 0.62$) and #1 ($MaxTS_{\#1} = 11.22$, $AvgTS_{\#1} = 3.12$). Participant #15 was the only participant to report regular use of virtual reality equipment, whereas

TABLE III: Within-subjects ANOVA on simulator sickness Total Score (TS)

Factor	Sum Sq	df	Mean Sq	F	p
VR experience	94627.08	3	31542.36	20.31	1.60E-11
Gender	11339.32	1	11339.32	5.81	0.0169 [†]
Texture resolution	1642.44	3	547.48	0.27	0.8463
Frame rate	1578.44	2	789.22	0.39	0.6759
Quality setting	2755.78	5	551.16	0.27	0.9287

[†] Biased due to only 2 out of 18 participants being female

TABLE IV: Average simulator sickness scores in relation to the subjects' prior VR experience

VR experience	N	O	D	TS
None	74.70	66.16	96.73	87.92
1-3 Times	40.54	43.51	55.68	52.33
Up to 10 / Regular	34.18	24.42	53.75	39.89

participant #1 reported having used virtual reality equipment 1-3 times. Both of them had prior experience with the HMD used in this study.

A within-subjects analysis of variance (ANOVA) was conducted in order to provide a statistical evaluation of the impact of prior experience with virtual reality equipment, gender, texture resolution, frame rate, and the quality setting (Resolution & Frame rate) on the participants' simulator sickness Total Scores (TS). The ANOVA results on Table III suggest that prior experience with virtual reality equipment had a statistically significant impact on simulator sickness TS ($p = 1.6 \cdot 10^{-11}$). Indeed, as shown on Table IV, average simulator sickness scores are the highest for subjects with no prior experience with VR, decrease for people that have used VR equipment 1-3 times, and are the lowest for subjects that reported having used VR equipment up to 10 times or more. Although gender was also shown to have a statistically significant impact ($p = 0.0169$), with female participants showing higher TS on average than male participants, this finding cannot be considered as reliable due to only 2 out of 18 participants being female. Moreover, results also indicated that resolution, frame rate, and quality setting did not have a statistically significant impact on simulator sickness TS ($p > 0.67$).

Furthermore, from Fig. 7 it is evident that the scores of disorientation (D) were always the highest among all examined settings, as also observed by Singla et al. [37] in a study about QoE and simulator sickness of omnidirectional videos in two commercially available HMDs, including the one used in this study. It is worth mentioning that in that study, the maximum average TS reported was less than 35, significantly less than the minimum average TS of 55 reported in this study. We attribute this difference to two main reasons: Firstly, the use of high resolution (FullHD and 4K) real video content with which the users are more familiar compared to the computer generated artificial virtual world used in this study,

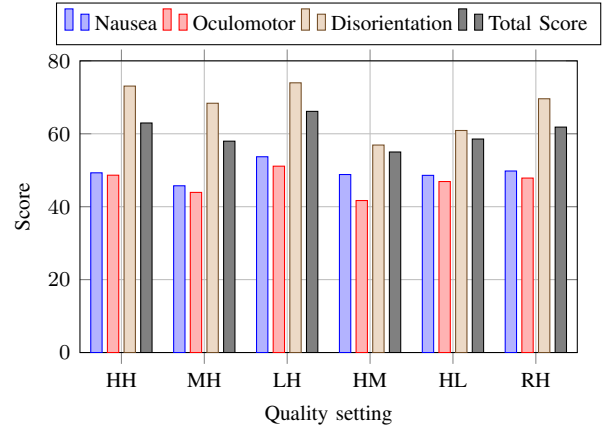


Fig. 7: Average simulator sickness scores for all the quality settings tested.

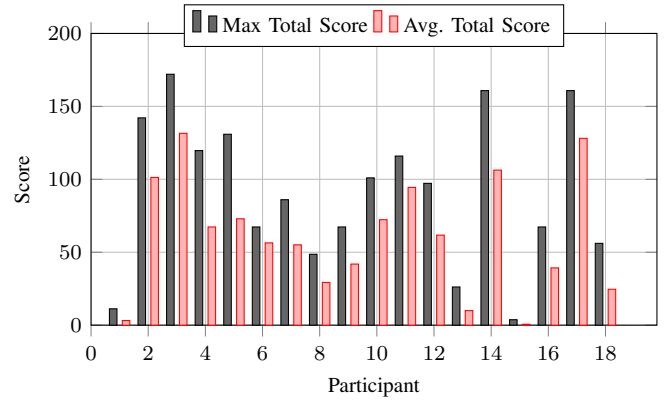


Fig. 8: Maximum and average simulator sickness Total Score for each participant during the experiment.

and secondly the fact that unlike being static observers of a video, in the examined application the users are engaged in constant body movement (bike pedalling) while at the same time they are exposed to more sudden changes in their viewing angle depending on their pedalling speed, steering, and course within the virtual world. As a result, all the examined quality settings included constant motion which exposed the users to constant sensory conflict due to the induced perception of self-motion caused by the optical flow patterns generated in the virtual environment. Sensory conflict is one of the main theories about the causes of simulator sickness [38], thus explaining the increased N, O, D, and TS scores in this study.

Out of the 18 participants of this study, 3 had to stop early due to severe symptoms of simulator sickness. All of them reported before the experiment that they had never used any virtual reality equipment. Participant #7 successfully completed 11 out of the 12 tests before quitting the experiment. During the session, his average simulator sickness Total Score (TS) was 55.08, with the maximum TS being 86.02. Participant #9 completed 5 tests, having an average TS of 41.89, with the maximum TS being 67.32, and participant #14 completed 7 tests, with a 106.32 average TS and a 160.82 maximum TS.

C. Physiological Signals Evaluation

The segments of the ECG and GSR recordings that corresponded to each test setting were extracted using the available timestamps, leading to a 60 *sec* ECG and GSR recording for each quality setting. Unfortunately, recordings for two participants were not retrievable and thus were not used for the analysis. Filtering was applied in order to reduce the levels of noise in the GSR signals. The filtering process consisted of a median filter with a time window of 500 *ms*, followed by a mean filter with a time window of 1000 *ms*. Lowpass filtering and smoothing are common pre-processing operations in GSR signal analysis [34], [39], employed to remove artefacts stemming from movement, as well as electrical artefacts.

After applying the filters, the number of peaks in each GSR segment was computed by detecting the local maxima with a minimum distance of 20 samples (≈ 1000 *ms*). Furthermore, the momentary heart rate (HR), corresponding to each sample of the ECG recording, was computed and five commonly used HR parameters [31], [32] were extracted from each ECG segment: the mean HR, minimum HR, maximum HR, median HR, and the standard deviation of the HR. An ANOVA was conducted in order to evaluate the impact of the quality settings on the aforementioned physiological parameters. The ANOVA results showed that the quality setting did not have a significant impact on the number of peaks in the GSR signal ($p = 0.2756$), neither on the examined HR parameters ($p > 0.8124$), as shown in Table V. Furthermore, a correlation analysis between the physiological parameters and the simulator sickness scores showed that there is only weak or very weak correlation between the two, as depicted in Table VI.

Sweating and increased heart rate are common symptoms associated with simulator sickness [20]. Nevertheless, the weak correlation found with the GSR and the heart rate parameters can be attributed to the specific activity performed in this experiment, i.e. pedalling on a bicycle. The sweating of both hands holding the bicycle's handlebars affected the GSR readings, whose electrodes were attached on the middle and index finger of the right hand of the participants. During strenuous activity, sweating will rise to dissipate the body heat generated, making it difficult to detect fluctuations in the GSR signal associated with emotional responses [34]. Similarly, the heart rate of the participants was also affected by the physical exercise conducted, thus explaining the weak correlation with simulator sickness scores. In terms of GSR and heart rate, the effects of simulator sickness did not produce responses of higher intensity compared to the physical exercise that the participants were engaged in, thus no connection could be established under the setting of the current experiment.

VI. CONCLUSION

In this work, we proposed and evaluated a smart-exercise-bike virtual reality system based on a commercially available HMD and a custom designed smart-bike controller. The proposed system, both as a generic exergaming controller, as well as a virtual reality exergaming system, constitutes a realistic and engaging solution for motivating users to engage in more physical exercise, thus providing health benefits to its

TABLE V: Within-subjects ANOVA between physiological parameters and quality ratings

Factor	Sum Sq	df	Mean Sq	F	p
GSR peaks	130.44	5	26.09	1.28	0.2756
Mean HR	159.47	5	31.89	0.25	0.9387
Min HR	146.97	5	29.39	0.19	0.9649
Max HR	316.62	5	63.32	0.42	0.8314
Median HR	173.03	5	34.61	0.27	0.9273
St.Dev. HR	8.52	5	1.70	0.45	0.8124

TABLE VI: Pearson's Correlation Coefficient between physiological parameters and simulator sickness scores

Parameter	TS	N	O	D
GSR peaks	-0.01	0.01	0.02	-0.05
Mean HR	-0.12	-0.17	-0.04	-0.15
Min HR	-0.09	-0.13	-0.00	-0.13
Max HR	-0.09	-0.13	-0.03	-0.12
Median HR	-0.13	-0.17	-0.04	-0.16
St.Dev. HR	0.06	0.07	0.01	0.09

users. In the case of the virtual reality application, the analysis of user feedback showed that texture quality and frame rate have a statistically significant impact on the perceived visual quality but not on the simulator sickness scores. Nevertheless, the moving speed of the virtual bike and the focus of the users into following the road and not crashing, prevented them from noticing any quality degradation when the frame rate was reduced from 60 fps to 30 fps. Furthermore, the quality setting was also shown to have no impact on the physiological responses of the participants in terms of the ECG and GSR modalities, while very weak correlation was observed between the physiological parameters and simulator sickness scores. Engaging in strenuous activity, like pedalling, causes an increase in sweating and in cardiac activity, thus interfering with physiological responses in the GSR and ECG signals that could be related to simulator sickness.

Increased simulator sickness scores were observed, compared to other works where the same HMD was used for viewing omnidirectional videos. This difference can be attributed to the use of computer generated graphics, as opposed to high resolution real video with which the users are more familiar, as well as to the constant and sometimes sudden motion that happens when riding the virtual bike, which results to increased sensory conflict. While the examined quality settings did not affect simulator sickness scores, the levels of prior experience with virtual reality technology were shown to have a statistically significant impact on simulator sickness Total Scores, with users more used to virtual reality technology being less affected by simulator sickness. Consequently, a slow familiarisation period could be allocated in order to allow users to become accustomed to the virtual reality system before engaging with it for long periods of time.

REFERENCES

- [1] O. Abari, D. Bharadia, A. Duffield, and D. Katabi, "Enabling high-quality untethered virtual reality," in *Proc. NSDI*, Boston, MA, USA, 2017, pp. 531–544.
- [2] M. Y. Boon, L. Asper, N. Jiao, and M. Ryan, "Vision training: comparing a novel virtual reality game of snakes with a conventional clinical therapy," in *Proc. IEEE LSC*, Sydney, Australia, 2017, pp. 43–46.
- [3] Y. Tanaka and M. Hirakawa, "Efficient strength training in a virtual world," in *Proc. IEEE ICCE-TW*, Nantou, Taiwan, 2016, pp. 1–2.
- [4] S. Shimada, K. Funahashi, K. Ito, Y. Tanase, and K. Iwazaki, "VR hand hygiene training system that visualizes germs to be washed and removed," in *Proc. IEEE GCCE*, Nagoya, Japan, 2017, pp. 1–4.
- [5] F. Buttussi and L. Chittaro, "Effects of different types of virtual reality display on presence and learning in a safety training scenario," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 2, pp. 1063–1076, Feb. 2018.
- [6] Z. Merchant, E. T. Goetz, L. Cifuentes, W. Keeney-Kennicutt, and T. J. Davis, "Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education," *Comput. Educ.*, vol. 70, no. C, pp. 29–40, Jan. 2014.
- [7] T. D. Parsons and J. L. Reinebold, "Adaptive virtual environments for neuropsychological assessment in serious games," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 197–204, May 2012.
- [8] E. Dzardanova, V. Kasapakis, and D. Gavalas, "On the effect of social context in virtual reality: An examination of the determinants of human behavior in shared immersive virtual environments," *IEEE Consum. Electron. Mag.*, vol. 7, no. 4, pp. 44–52, Jul. 2018.
- [9] S. C. Barathi, D. J. Finnegan, M. Farrow, A. Whaley, P. Heath, J. Buckley, P. W. Dowrick, B. C. Wuensche, J. L. J. Bilzon, E. O'Neill, and C. Lutteroth, "Interactive feedforward for improving performance and maintaining intrinsic motivation in VR exergaming," in *Proc. ACM CHI*, Montreal, Canada, 2018, pp. 408:1–408:14.
- [10] J. Sinclair, P. Hingston, M. Masek, and K. Nosaka, "Testing an exergame for effectiveness and attractiveness," in *Proc. IEEE ICE-GIC*, Hong Kong, China, 2010, pp. 1–8.
- [11] J. Sinclair, P. Hingston, and M. Masek, "Considerations for the design of exergames," in *Proc. ACM GRAPHITE*, Perth, Australia, 2007, pp. 289–295.
- [12] L. Müller, S. Zagaria, A. Bernin, A. Amira, N. Ramzan, C. Grecos, and F. Vogt, "Emotionbike: A study of provoking emotions in cycling exergames," in *Proc. IFIP ICEC*, Trondheim, Norway, 2015, pp. 155–168.
- [13] A. Laikari, "Exergaming - Gaming for health: A bridge between real world and virtual communities," in *Proc. IEEE ISCE*, Kyoto, Japan, 2009, pp. 665–668.
- [14] I.-M. Lee, E. J. Shiroma, F. Lobelo, P. Puska, S. N. Blair, and P. T. Katzmarzyk, "Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy," *The Lancet*, vol. 380, no. 9838, pp. 219–229, Jul. 2012.
- [15] G. Williams and G. Frühbeck, *Obesity: Science to practice*. Hoboken, NJ, USA: Wiley, 2009.
- [16] American College of Sports Medicine, *ACSM's guidelines for exercise testing and prescription*, 10th ed. Philadelphia, PA, USA: Wolters Kluwer, 2017.
- [17] J. Annesi, "Effects of a cognitive behavioral treatment package on exercise attendance and drop out in fitness centers," *Eur. J. Sport Sci.*, vol. 3, no. 2, pp. 1–16, Nov. 2003.
- [18] K. Hagen, K. Chorianopoulos, A. I. Wang, L. Jaccheri, and S. Weie, "Gameplay as exercise," in *Proc. ACM CHI*, San Jose, CA, USA, 2016, pp. 1872–1878.
- [19] G.-A. Koulouris, B. Bui, M. S. Banks, and G. Drettakis, "Accommodation and comfort in head-mounted displays," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 87:1–87:11, Jul. 2017.
- [20] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness," *Int. J. Aviat. Psychol.*, vol. 3, no. 3, pp. 203–220, 1993.
- [21] P. Arnau-González, T. Althobaiti, S. Katsigiannis, and N. Ramzan, "Perceptual video quality evaluation by means of physiological signals," in *Proc. QoMEX*, Erfurt, Germany, 2017, pp. 1–6.
- [22] J. Bakker, M. Pechenizkiy, and N. Sidorova, "What's your current stress level? Detection of stress patterns from GSR sensor data," in *Proc. IEEE ICDMW*, Vancouver, Canada, 2011, pp. 573–580.
- [23] H. Heo, E. C. Lee, K. R. Park, C. J. Kim, and M. Whang, "A realistic game system using multi-modal user interfaces," *IEEE Trans. Consum. Electron.*, vol. 56, no. 3, pp. 1364–1372, Aug. 2010.
- [24] D. E. Warburton, S. S. Bredin, L. T. Horita, D. Zbogor, J. M. Scott, B. T. Esch, and R. E. Rhodes, "The health benefits of interactive video game exercise," *Appl. Physiol. Nutr. Metab.*, vol. 32, no. 4, pp. 655–663, Aug. 2007.
- [25] J. R. Bruun-Pedersen, S. Serafin, and L. B. Kofoed, "Motivating elderly to exercise - recreational virtual environment for indoor biking," in *Proc. IEEE SeGAH*, Orlando, FL, USA, 2016, pp. 1–9.
- [26] S. Göbel, S. Hardy, V. Wendel, F. Mehm, and R. Steinmetz, "Serious games for health: Personalized exergames," in *Proc. ACM MM*, Firenze, Italy, 2010, pp. 1663–1666.
- [27] M. Hoda, R. Alattas, and A. E. Saddik, "Evaluating player experience in cycling exergames," in *Proc. IEEE ISM*, Anaheim, CA, USA, 2013, pp. 415–420.
- [28] J. M. Silva and A. El Saddik, "Exertion interfaces for computer videogames using smartphones as input controllers," *Multimedia Syst.*, vol. 19, no. 3, pp. 289–302, Jun. 2013.
- [29] F. Grani and J. R. Bruun-Pedersen, "Giro: Better biking in virtual reality," in *Proc. IEEE WEVR*, Los Angeles, CA, USA, 2017, pp. 1–5.
- [30] A. Burns, B. R. Greene, M. J. McGrath, T. J. O'Shea, B. Kuris, S. M. Ayer, F. Stroiescu, and V. Cionca, "Shimmer™ - A wireless sensor platform for noninvasive biomedical research," *IEEE Sensors J.*, vol. 10, no. 9, pp. 1527–1534, Sept. 2010.
- [31] S. Katsigiannis and N. Ramzan, "DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 1, pp. 98–107, Jan. 2018.
- [32] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan. 2012.
- [33] Cooking Hacks, "e-Health sensor platform v2.0 for Arduino and Raspberry Pi [biometric / medical applications]," accessed 23 January 2018. [Online]. Available: <https://www.cooking-hacks.com/documentation/tutorials/ehealth-biometric-sensor-platform-arduino-raspberry-pi-medical/>
- [34] W. Boucsein, D. C. Fowles, S. Grimnes, G. Ben-Shakhar, W. T. Roth, M. E. Dawson, D. L. Fillion, W. Boucsein, D. C. Fowles, S. Grimnes, G. Ben-Shakhar, W. T. Roth, M. E. Dawson, and D. L. Fillion, "Publication recommendations for electrodermal measurements," *Psychophysiology*, vol. 49, no. 8, pp. 1017–1034, Aug. 2012.
- [35] ITU-R, "Subjective video quality assessment methods for multimedia applications," ITU-R Rec. P.910, Apr. 2008.
- [36] R. S. Kellogg, R. S. Kennedy, and A. Graybiel, "Motion sickness symptomatology of labyrinthine defective and normal subjects during zero gravity maneuvers," *Aerosp. Med.*, vol. 36, pp. 315–318, Apr. 1965.
- [37] A. Singla, S. Fremerey, W. Robitz, and A. Raake, "Measuring and comparing QoE and simulator sickness of omnidirectional videos in different head mounted displays," in *Proc. QoMEX*, Erfurt, Germany, 2017, pp. 1–6.
- [38] D. M. Johnson, "Introduction to and review of simulator sickness research," U.S. Army Research Institute for the Behavioral and Social Sciences, Tech. Rep. 1832, Apr. 2005.
- [39] J. Braithwaite, D. Watson, R. Jones, and M. Rowe, "A guide for analysing electrodermal activity (EDA) and skin conductance responses (SCRs) for psychophysiological experiments," University of Birmingham, Selective Attention & Awareness Laboratory (SAAL), Tech. Rep. 2nd version, 2015.